

Inter-Domain Routing
Internet-Draft
Updates: 4271 (if approved)
Intended status: Standards Track
Expires: 7 January 2026

Z. Li, Ed.
L. Song, Ed.
China Mobile
6 July 2025

BGP SR Policy Extensions for Performance-Aware Path Selection
draft-li-idr-sr-policy-metric-01

Abstract

To enable the headend node to do performance-aware path selection, this document proposes an extension to the BGP SR Policy protocol by defining a new optional Metric Sub-TLV within the BGP Tunnel Encapsulation Attribute [RFC9012]. The introduced Metric Sub-TLV encodes performance parameters (such as latency, bandwidth, reliability, etc.) for SR Policy paths.

This specification also updates the BGP route selection procedures in [RFC4271], modifying the Breaking Ties (Phase 2) logic to prioritize the metrics for SR Policy paths.

Key contributions include:

- * Introduce Metric Sub-TLV in BGP SR Policy
- * Update the tie-breaking procedure for BGP route selection

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 7 January 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
1.1. Requirements Language	3
2. Use Case	3
3. Metric Sub-TLV	4
4. Policy Computation and Provisioning System Behavior	6
5. Headend Behavior	7
6. Updated Tie-Breaking Procedure for BGP	7
7. IANA Considerations	8
8. Security Considerations	8
9. References	8
9.1. Normative References	8
9.2. Informative References	9
Acknowledgements	9
Authors' Addresses	9

1. Introduction

Segment Routing (SR) [RFC8402] allows a headend node to steer a packet flow along a specific path. [RFC9256] further details the concepts of SR Policy and steering into an SR Policy. [I-D.ietf-idr-sr-policy-safi] specifies the use of BGP to distribute one or more of the candidate paths of an SR Policy to the headend of that policy. Currently [I-D.ietf-idr-sr-policy-safi] lacks the capability to propagate performance metrics such as path latency, bandwidth, or reliability. This limitation prevents headends from implementing policy selection based on path metrics when there are multiple paths reaching the same destination. Consequently, the headends cannot dynamically elect performance-optimal path among multiple SR Policies. To address this limitation, this document extends the BGP SR Policy protocol to carry performance metrics in SR Policy advertisements.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Use Case

As shown in Figure 1, the SR Policy Computation and Provisioning System, such as a SDN controller, collects real-time network state information (e.g., topology, link utilization) and performance metrics (e.g., link latency, jitter, packet loss rate). Based on service or customer requirements (e.g., minimum latency), it computes SR Policy paths between designated endpoints and distributes them to headend nodes via the BGP SR Policy protocol [I-D.ietf-idr-sr-policy-safi].

For example:

The system provisions two low-latency policies to headend node PE3:

Policy1: Path via P1-->PE1, with a measured latency of 20 ms.

Policy2: Path via P2-->PE2, with a measured latency of 12 ms.

However, the current BGP SR Policy protocol [I-D.ietf-idr-sr-policy-safi] only propagates path definitions (e.g., segment lists) without embedding performance metrics. This forces headend nodes to select paths based solely on static criteria (e.g., administrative preferences), potentially leading to suboptimal traffic engineering decisions.

To address this limitation, this proposal extends the BGP SR Policy protocol by introducing a new Performance Metrics Sub-TLV within the BGP Tunnel Encapsulation Attribute [RFC9012]. This Sub-TLV encodes key performance indicators (KPIs) such as latency, bandwidth, and reliability (see Section 3 for details). With this extension: The SR Policy Computation and Provisioning System can advertise SR Policies alongside their associated KPIs. Headend nodes leverage the enhanced BGP route selection logic (Section 6) to prioritize paths that meet dynamic performance requirements.

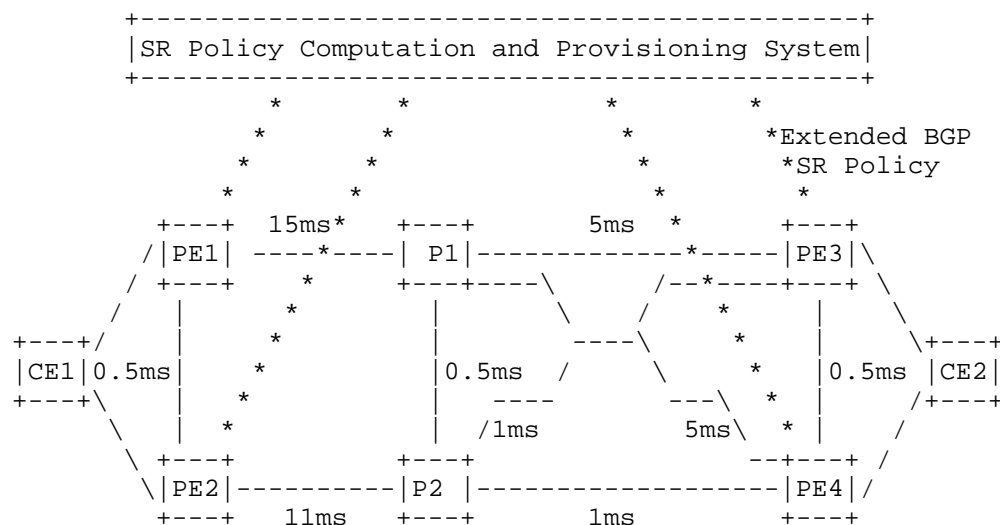


Figure 1: Use Case for Performance-Aware SR Policy Selection

3. Metric Sub-TLV

This document extends the BGP SR Policy protocol by introducing a new sub-TLV, Metric Sub-TLV, within the BGP Tunnel Encapsulation Attribute. The Extended BGP SR Policy Encoding structure is as follows:

SR Policy SAFI NLR1: <Distinguisher, Policy-Color, Endpoint>
Attributes:

```
Tunnel Encapsulation Attribute(23)
  Tunnel Type: SR Policy(15)
    Binding SID
    SRv6 Binding SID
    Preference
    Priority
    Metric
    Policy Name
    Policy Candidate Path Name
    Explicit NULL Label Policy (ENLP)
    Segment List
      Weight
      Segment
      Segment
      ...
    ...
```

Figure 2: Extended BGP SR Policy Encoding

Metric Sub-TLV is used to carry performance metrics such as latency, bandwidth, and reliability. The format of Metric Sub-TLV is as follows:

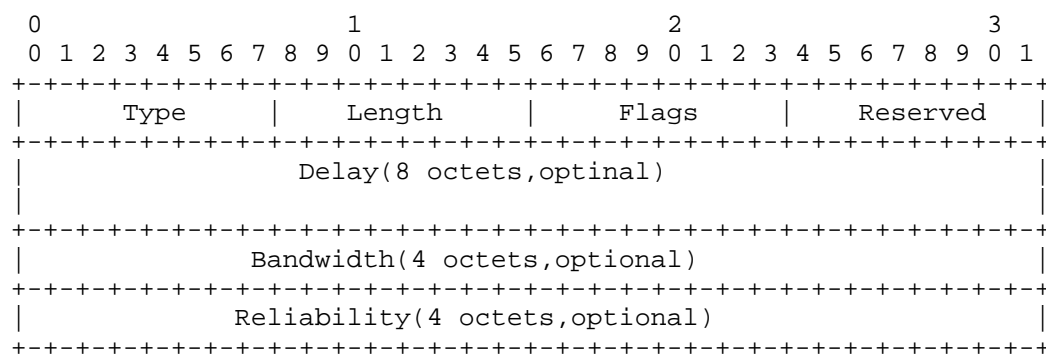


Figure 3: Metric Sub-TLV

Where:

Type (1 octet): Indicates this sub-TLV is Metric, Specific values need to be assigned by IANA.

Length (1 octet): Indicates the length of the Metric sub-TLV in bytes.

Flags (1 octet): Indicates the presence of specific performance metrics. Its definition is shown in Figure 4.

Reserved (1 octet): Reserved for future use. This field MUST be set to 0 when sending and ignored when receiving.

Delay(8 octets): Carries delay information. Its format depends on the D flag in Flags Field:

- * If D = 01: NTPv4 format delay
- * If D = 10: IEEE 1588v2 PTP format delay

Bandwidth (4 octets): Carries bandwidth information in Mbps.

Reliability (4 octets): Carries reliability information, such as the maximum number of failures that have occurred on all links in the SR policy path within the past month.

```

      0 1 2 3 4 5 6 7
    +---+---+---+---+
    | D |B|R|Reserved|
    +---+---+---+---+

```

Figure 4: Flag Field for Metric Sub-TLV

Where, D Flag is for delay, B Flag is for Bandwidth and R Flag is for reliability, all the other bits are reserved. The detailed encodings of the three flags defined in this document are as follows:

Flag	Bits	Description
D	0-1	00: No delay 01: NTPv4 delay 10: PTP delay 11: Reserved
B	2	0: No bandwidth 1: Bandwidth
R	3	0: No reliability 1: Reliability

Figure 5: Flags for Metric Sub-TLV

Implementations SHOULD set only one flag (D, B, or R) at a time, as these metrics are typically not directly comparable. Network operators MAY configure which metric to prioritize based on service requirements.

4. Policy Computation and Provisioning System Behavior

The Policy Computation and Provisioning System is responsible for calculating Segment Routing (SR) policies based on network state and business requirements, and provisioning them to headend nodes. When provisioning SR policies that include performance metrics, the system should follow these steps:

Collect Network State Information: Gather real-time network topology, link utilization, and other relevant data.

Compute SR Policies: Calculate SR policy paths that meet performance requirements based on service needs and network state.

Encapsulate Performance Metrics: Embed performance metrics such as latency, bandwidth, and reliability within the Metric Sub-TLV of the BGP Tunnel Encapsulation Attribute.

Provision BGP Update Messages: Include the SR policies with performance metrics in BGP update messages and send them to the appropriate headend nodes.

5. Headend Behavior

Upon receiving SR policies with performance metrics, headend nodes should process them as follows:

Parse BGP Update Messages: Extract SR policies and their associated performance metrics from the received BGP update messages.

Store Performance Metrics: Save the performance metrics in a local database for subsequent path selection.

Path Selection: Prioritize paths that meet dynamic performance requirements when multiple paths are available.

Update Routing Tables: Modify routing tables based on the selected paths to ensure traffic is forwarded along optimized routes.

6. Updated Tie-Breaking Procedure for BGP

Support for SR policy metric introduced in this document involves several modifications to the tie-breaking procedures of the BGP "phase 2" decision described in [RFC4271], Section 9.1.2.2.

A new step, step e0, is inserted before step e of the tie-breaking (Phase 2) logic in [RFC4271].

e0) If any routes have the Color extended community attribute with an identical value, remove from consideration all routes lacking the Color extended community attribute.

Compare the SR policies corresponding to the remaining routes. If any SR policies have the metric sub-TLV, remove from consideration all routes whose corresponding SR policies do not have the metric sub-TLV.

Remove from consideration all routes whose metric sub-TLV are not the best. For the latency metric, the smallest value is considered the best; for the bandwidth metric, the largest value is the best; and for reliability, the smallest value is the best.

7. IANA Considerations

IANA is requested to assign the following code point from the "BGP Tunnel Encapsulation Attribute Sub-TLVs" Registry:

Code Point	Description	Reference
TBD	Metric Sub-TLV	This document

Table 1: Code Point for Metric Sub-TLV

8. Security Considerations

TBD

9. References

9.1. Normative References

- [I-D.ietf-idr-sr-policy-safi]
Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-sr-policy-safi-13, 6 February 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-sr-policy-safi-13>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <<https://www.rfc-editor.org/info/rfc4271>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <<https://www.rfc-editor.org/info/rfc9012>>.

9.2. Informative References

- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <<https://www.rfc-editor.org/info/rfc9256>>.

Acknowledgements

The authors would like to acknowledge the supports from Cheng Chang and Bo Liu.

Yao Liu and Changwang Lin review this document and provide valuable comments.

Authors' Addresses

Zhenqiang Li (editor)
China Mobile
China
Email: lizhenqiang@chinamobile.com

Liyan Song (editor)
China Mobile
China
Email: songliyan@chinamobile.com